

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2000-259719

(43)Date of publication of application : 22.09.2000

(51)Int.Cl.

G06F 17/60

G06F 9/44

G06F 15/18

(21)Application number : 11-059885

(71)Applicant : INTERNATL BUSINESS MACH CORP
<IBM>

(22)Date of filing : 08.03.1999

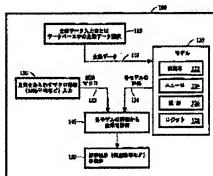
(72)Inventor : MORIMOTO YASUHIKO
KUMADA KIYOSHI
FUKUDA TSUYOSHI

(54) METHOD AND DEVICE FOR CALCULATING PROBABILITY OF DEFAULT ON OBLIGATION

(57)Abstract:

PROBLEM TO BE SOLVED: To provide a method for calculating the probability of a bankruptcy which is capable of calculating actual probability regardless of the ratio of bankruptcies/non-bankruptcies in learned data.

SOLUTION: The method for calculating the probability of the occurrence of a default on an obligation of a company includes (a) a step for inputting the financial data of a plurality of companies, (b) a step for preparing a decision tree from the financial data, (c) a step for applying the financial data of an objective company to the decision tree and (d) a step for calculating the probability of the occurrence of the default on an obligation from the result of applying the decision tree to the objective company. In addition, (d) the step utilizes the Bayes' theorem.



(51)Int.Cl. ⁷	識別記号	F I	テーマコード(参考)
G 0 6 F 17/60		G 0 6 F 15/21	Z 5 B 0 4 9
9/44	5 5 0	9/44	5 5 0 N
15/18	5 5 0	15/18	5 5 0 A

審査請求 有 請求項の数17 O L (全 14 頁)

(21)出願番号 特願平11-59885

(22)出願日 平成11年3月8日(1999.3.8)

(71)出願人 390009531

インターナショナル・ビジネス・マシーンズ・コーポレーション

INTERNATIONAL BUSINESS MACHINES CORPORATION

アメリカ合衆国10504、ニューヨーク州

アーモンク (番地なし)

(74)代理人 100086243

弁理士 坂口 博 (外1名)

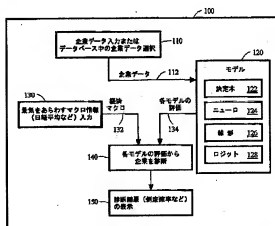
最終頁に続く

(54)【発明の名称】 債務不履行の確率を計算する方法及び装置

(57)【要約】

【課題】本発明の課題は、学習データ内の倒産/非倒産の比率によらず現実的な確率を計算することができる倒産確率の計算方法を提供することである。

【解決手段】本発明の構成は、企業の債務不履行の発生確率を計算する方法が、(a)複数の企業の財務データを入力するステップ、(b)前記財務データから決定木を作成するステップ、(c)対象企業の財務データを決定木に適用するステップ、(d)対象企業を決定木に適用した結果から債務不履行の発生確率を計算するステップを含む。更に、前記(d)ステップが、ベイズの定理を利用することの特徴とする。



【特許請求の範囲】

【請求項1】企業の債務不履行の発生確率を計算する方法であって、

前記方法は、(a)複数の企業の財務データを入力するステップと、(b)前記財務データから決定木を作成するステップと、(c)対象企業の財務データを決定木に適用するステップと、(d)前記対象企業を決定木に適用した結果から債務不履行の発生確率を計算するステップと、

を含む方法。

【請求項2】企業の債務不履行の発生確率を計算する方法であって、

前記方法は、(a)複数の企業の財務データを入力するステップと、(b)前記財務データから決定木を作成するステップと、(c)対象企業の財務データを決定木に適用するステップと、(d)景気を表すデータから平均的倒産率を計算するステップと、(e)前記対象企業を決定木に適用した結果及び前記平均的倒産率から債務不履行の発生確率を計算するステップと、

を含む方法。

【請求項3】企業の債務不履行の発生確率を計算するプログラムを記録した記録媒体であって、

前記プログラムは、(a)複数の企業の財務データを入力するステップと、(b)前記財務データから決定木を作成するステップと、(c)対象企業の財務データを決定木に適用するステップと、(d)前記対象企業を決定木に適用した結果から債務不履行の発生確率を計算するステップと、

を含むプログラムを記録した記録媒体。

【請求項4】企業の債務不履行の発生確率を計算するプログラムを記録した記録媒体であって、

前記プログラムは、(a)複数の企業の財務データを入力するステップと、(b)前記財務データから決定木を作成するステップと、(c)対象企業の財務データを決定木に適用するステップと、(d)景気を表すデータから平均的倒産率を計算するステップと、(e)前記対象企業を決定木に適用した結果及び前記平均的倒産率から債務不履行の発生確率を計算するステップと、

を含むプログラムを記録した記録媒体。

【請求項5】前記(d)ステップが、ベイズの定理を利用することを特徴とする請求項1乃至4に記載の方法。

【請求項6】企業の債務不履行の発生確率を計算する装置であって、

前記装置は、(a)複数の企業の財務データを入力する手段と、(b)前記財務データから決定木を作成する手段と、(c)対象企業の財務データを決定木に適用する手段と、(d)前記対象企業を決定木に適用した結果から債務不履行の発生確率を計算する手段と、

を含む装置。

【請求項7】企業の債務不履行の発生確率を計算する装

置であって、

前記装置は、(a)複数の企業の財務データを入力する手段と、(b)前記財務データから決定木を作成する手段と、(c)対象企業の財務データを決定木に適用する手段と、(d)景気を表すデータから平均的倒産率を計算する手段と、(e)前記対象企業を決定木に適用した結果及び前記平均的倒産率から債務不履行の発生確率を計算する手段と、

を含む装置。

10 【請求項8】企業の債務不履行の発生確率を計算するシステムであって、

前記システムは、プロセッサ、記憶装置、入力装置、表示装置を含み、更に、(a)複数の企業の財務データを入力する手段と、(b)前記財務データから決定木を作成する手段と、(c)対象企業の財務データを決定木に適用する手段と、(d)前記対象企業を決定木に適用した結果から債務不履行の発生確率を計算する手段と、

を含むシステム。

20 【請求項9】企業の債務不履行の発生確率を計算するシステムであって、

前記システムは、プロセッサ、記憶装置、入力装置、表示装置を含み、更に、(a)複数の企業の財務データを入力する手段と、(b)前記財務データから決定木を作成する手段と、(c)対象企業の財務データを決定木に適用する手段と、(d)景気を表すデータから平均的倒産率を計算する手段と、(e)前記対象企業を決定木に適用した結果及び前記平均的倒産率から債務不履行の発生確率を計算する手段と、

を含むシステム。

30 【請求項10】前記(d)手段が、ベイズの定理を利用することを特徴とする請求項6乃至9に記載の方法。

【請求項11】前記財務データが、倒産企業のデータ及び非倒産企業のデータを含むことを特徴とする請求項1乃至10に記載の方法。

【請求項12】自動車事故の発生確率を計算する方法であって、

前記方法は、(a)複数の自動車事故データを入力するステップと、(b)前記自動車事故データから決定木を作成するステップと、(c)対象となる自動車事故データを決定木に適用するステップと、(d)前記対象となる自動車事故データを決定木に適用した結果から対象となる自動車事故の発生確率を計算するステップと、

を含む方法。

【請求項13】火災の発生確率を計算する方法であって、

前記方法は、(a)複数の火災データを入力するステップと、(b)前記火災データから決定木を作成するステップと、(c)対象となる火災データを決定木に適用するステップと、(d)前記対象となる火災データを決定木に適用した結果から対象となる火災の発生確率を計算

するステップと、を含む方法。

【請求項 14】企業の債務不履行の発生確率を計算する方法であって、

前記方法は、(a)複数の企業の財務データを入力するステップと、(b)前記財務データから決定木を作成するステップと、(c)対象企業の財務データを決定木に適用するステップと、(d)前記対象企業を決定木に適用した結果から債務不履行の発生確率を計算するステップを含む、

前記(c)ステップが、前記複数の企業を倒産企業が多いグループと非倒産企業が多いグループの2つのグループに分類するためのデータを前記財務データの中から発見するステップを含む方法。

【請求項 15】人の死亡確率を計算する方法であって、前記方法は、

(a)複数の人のデータを入力するステップと、(b)前記人のデータから決定木を作成するステップと、(c)対象となる人のデータを決定木に適用するステップと、(d)前記対象となる人のデータを決定木に適用した結果から対象となる人の死亡確率を計算するステップと、を含む方法。

【請求項 16】地震の発生確率を計算する方法であって、

前記方法は、(a)複数の地震のデータを入力するステップと、(b)前記地震のデータから決定木を作成するステップと、(c)対象となる地震のデータを決定木に適用するステップと、(d)前記対象となる地震のデータを決定木に適用した結果から対象となる地震の発生確率を計算するステップと、を含む方法。

【請求項 17】前記(d)ステップが、ベイズの定理を利用することを特徴とする請求項 12乃至16に記載の方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、対象企業の債務不履行の確率を計算可能なコンピュータ・システム、特に決定木を利用した債務不履行の確率を計算する方法及び装置に関する。

【0002】

【従来の技術】近年、データベースから属性やデータ間に成り立つ法則・ルールを発掘し、経営戦略を立案する上で役に立つ法則を得る「データ・マイニング」と呼ばれる研究が盛んに行われている。

【0003】その成果はあらゆる業界のデータ分析に応用され、特に金融業界は、データ分析能力を、経営の死命を決することもある重要な技術として認識し、古くから多変量解析、シミュレーションなど様々なデータ分析技術を応用してきた。

【0004】金融機関はいまでもなく、あらゆる企業の財務部門や審査部門においては、取引先の企業の「倒産及び、それにとまなう債務不履行」の危険性がどの程度あるのか見定めることは、その企業の経営を左右するとともに非常に困難な課題である。

【0005】企業の倒産の危険性を数値として算出する、いわゆる、信用リスクの定量化システムは、今後ますますその重要性を増していくと考えられる。

【0006】本発明は、そうした重要なニーズに応じるべく、企業の危険度の1つの目安として財務データから「倒産確率(又は債務不履行確率)」の数値を算出するための手法である。

【0007】企業の過去の財務データから、倒産、非倒産を判定する決定木を作成することができる。決定木は、企業の集合を財務データの内容から繰り返し、ある代表的な(相互情報量などを最適化するような)ルールで部分集合に分割してゆくことで構成される。

【0008】こうした分割により、最終的に作られるそれぞれの部分集合は「倒産」又は「非倒産」のいずれかに分布が偏ったものとなる。

【0009】審査対象企業はこのような作られた決定木により、いずれかの部分集合の1つとして分類され、その部分集合が持っていた倒産の頻度から「倒産」か「非倒産」かを判定される。

【0010】決定木は、判定結果を算出するとともに、その判定に至る根拠をルールとして示すことができるため、分析ツールとして広く用いられている。

【0011】しかし、信用リスクの定量化という観点からは「倒産である」又は「非倒産である」といった離散的で定性的な判定結果を出力するより、どのくらいの確率で倒産するか、倒産しないかといった連続数値を定量的に出力する方が、この種のアプリケーションとしての意義が大きいと考えられる。

【0012】判別を目的に作成される決定木から、判別結果ではなく、特定の判別に至る確率を求めるには、学習データ(サンプル)と母集団の関係等を注意深く考慮する必要がある。

【0013】更に、倒産という事象に注目するならば、正確な確率を求めるためには、現在の景気の動向を簡単に反映させることができないとなければならない。

【0014】本発明は、決定木を使って債務不履行(すなわち倒産)の確率を算出するというアプリケーションに特化した問題解決法を提供するものである。

【0015】決定木は、判定結果を算出するとともに、その判定に至る根拠をルールとして示すことができるため、分析ツールとして広く用いられている。しかしながら、信用リスクの定量化という観点からは「倒産である」、「非倒産である」といった離散的な判定結果を出力するより、どのくらいの確率で倒産するか、倒産しないかといった連続数値として出力する方が、このアプリ

ケーションとしての意義が大きい。

【0016】決定木の情報から倒産確率のような数値を求めるには、以下のような問題点がある。

【0017】すなわち、現実には全企業数に比べ倒産の頻度は極めて小さいため、倒産企業の特徴、非倒産企業の特徴をそれぞれ学習するためには、倒産する企業数も十分な数が必要になる。

【0018】そのため現実には比べ倒産企業の多いデータ集合を学習データとして使うことになり、決定木で作られる部分集合内の倒産頻度自体は現実的な数値とはいえない。

【0019】また、企業倒産は、景気状況などマクロ経済的な影響にも左右されやすいので、学習データから得た倒産頻度を単純に現実的な数値に置き換えるだけでは不十分である。

【0020】決定木で、学習データから作られる部分集合における倒産/非倒産の分布比は、その部分集合に至る条件を満たした場合の条件付き確率であると読み替えることができる。

【0021】ある企業が、ある部分集合に分類された場合、ベイズの理論を利用して、この条件付き確率をもとに、倒産という事象の（この部分集合に分類されたという）事後確率を求めることができる。

【0022】また、事後確率に対する事前確率、つまり、どこにも分類されていない状態での（財務状況を見ていない状態での）企業の倒産確率は、景気の状態を反映した確率とする。

【0023】景気の状態を反映した指標はいろいろ存在するが、誰もが手軽に知ることができ、かつ、企業倒産数と相関が強い指標の1つに日経平均株価がある。

【0024】日経平均株価から事前確率である全企業の倒産確率を算定する。これにより、ベイズの事後確率に相当する各企業の倒産確率に、景気の動向を反映させることができる。

【0025】

【発明が解決しようとする課題】本発明の目的は、学習データ内の倒産/非倒産の比率によらず現実的な確率を計算することができる倒産確率の計算方法を提供することである。

【0026】また、本発明の他の目的は、学習データ内の倒産/非倒産の比率によらず現実的な確率を計算することができる、かつ、計算時点の景気状況を確率に反映させることができる倒産確率の計算方法を提供することである。

【0027】

【課題を解決するための手段】本発明の第1の実施例では、企業の債務不履行の発生確率を計算する方法が、

(a) 複数の企業の財務データを入力するステップと、

(b) 前記財務データから決定木を作成するステップと、

(c) 対象企業の財務データを決定木に適用するス

テップと、(d) 対象企業を決定木に適用した結果から債務不履行の発生確率を計算するステップとを含む。

【0028】本発明の第2の実施例では、企業の債務不履行の発生確率を計算する方法が、(a) 複数の企業の財務データを入力するステップと、(b) 前記財務データから決定木を作成するステップと、(c) 対象企業の財務データを決定木に適用するステップと、(d) 景気を表すデータから平均的倒産率を計算するステップと、(e) 対象企業を決定木に適用した結果及び前記平均的倒産率から債務不履行の発生確率を計算するステップと、を含む。

【0029】本発明の第3の実施例では、企業の債務不履行の発生確率を計算する方法が、(a) 複数の企業の財務データを入力するステップと、(b) 前記財務データから決定木を作成するステップと、(c) 対象企業の財務データを決定木に適用するステップと、(d) 景気を表すデータから平均的倒産率を計算するステップと、(e) 対象企業を決定木に適用した結果及び前記平均的倒産率から債務不履行の発生確率を計算するステップと、を含む、前記(d)ステップが、ベイズの定理を利用する。

【0030】

【発明の実施の形態】図1には、本発明を適用した企業の倒産確率計算プログラムの機能的構成が示されている。企業の倒産確率計算プログラムは、主たる機能的な構成要素として、企業データの入力・選択部110、倒産確率を計算するためのモデル部120、景気を表すマクロ情報の入力部130、各モデルの評価から企業を診断する診断部140、及び診断結果の表示部150を含んでいる。

【0031】企業データの入力・選択部110は、企業の財務データを読み込む。財務データの読み込み方法は、データベースへの定期的な一括入力や、画面からの逐次入力等がある。財務データとしては、例えば帝国データバンク社のものがある。この財務データは、社名ファイル、格付ファイル、財務ファイル等を含んでいる。

【0032】また、財務ファイルは、各企業の決算期年、決算期月、従業員数、中分類業種及びB/S、P/L及び利益処分計算書などのデータが含まれるのが一般的である。企業データの入力・選択部110は、次に、入力された財務データから財務指標を算出する。

【0033】倒産確率を計算するためのモデル部120は、企業データの入力・選択部110から企業データ、財務指標等のデータを受け取り各モデルを用いて評価を計算する。倒産確率を計算するためのモデル部120には、例えば、決定木モデル122、ニューロ・モデル124、線形モデル126、ロジック・モデル128等が含まれる。

【0034】ここで、モデルとは、最適化されたアルゴリズムのことである。モデルは、企業の信用力を財務デ

ータから評価し、「要注意」な企業なのか、「安全」な企業なのかを判別する。

【0035】ニューロ・モデル124とは、ニューロ判別を最適化するように作られたモデルであり、線形モデル126とは、財務指標を統計的に選択して最適化された線形判別モデルであり、ロジック・モデル128とは、財務指標を統計的に選択して最適化されたモデルである。

【0036】決定木モデル122は、データマイニングという技術を用いたモデルである。データマイニングとは、大量なデータの集まり（データウェアハウジング）から有用な情報を引き出すための技術であり、「データベースからの知識獲得」とも呼ばれている。

【0037】これは、大量のデータが集まれば、そのデータを解析することで、それまで気づかなかった新たな知識が得られるというものである。

【0038】例えば、医療診断、POSデータからの購買パターンの解析、クレジットカードの使用者の信用度、通信販売での顧客管理など、電子的な情報の蓄積が進むとともに多くの分野に応用されてきている。

【0039】決定木は、このデータマイニングの手法の1つであり、複数の数値属性を持つ大量のデータを2つのグループに分類する技術である。例えば、診察データから患者が病気の可能性が高い/低いという2つのグループに分類したり、財務データから企業が倒産しやすい/しにくいという2つのグループに分類することが可能となる。

【0040】決定木は、こうした分類を2分木構造で実現する。それぞれの分岐点はノードと呼ばれ、木の根（ルート）から出発し、ノード毎に左右いずれかを選びつつ下っていき、最終的に到達したノードが対応する分類結果として採用される。

【0041】決定木の性能としては、分類性が高いこと、すなわち、最終ノードに到達するまでに如何にうまく2つのグループを切り分けることができるかという点が重要である。また、最終的に完成した木構造がなるべく単純であることが望ましく、木の深さ（階層）は浅く、全ノードの数が少ない程良いモデルということになる。

【0042】2つの子孫を持つ各ノードは、データに対するテストに対応する。例えば、自己資本率がある値（閾値）より大きいか、小さいかによって枝を左右のどちらに進むかが決定される。このようなテストにどの指標又はデータを使うか、或いは、閾値をどう設定するかが決定木の本質となる。

【0043】景気を表すマクロ情報の入力部130から、景気を表すマクロ情報、例えば日経平均株価等のデータとその日付を入力する。

【0044】各モデルの評価から企業を診断する部140は、各モデル120からの各モデルの評価結果及び景

気を表すマクロ情報の入力部130から経済マクロ情報を得て、対象企業を診断する。

【0045】診断結果の表示部150は、各モデルの評価から企業を診断する部140から得た評価結果をユーザに提示する。

【0046】図2は、本発明を適用したシステムの主な処理の流れを示したフローチャートである。ステップ210において、本発明を適用したシステムが開始される。

【0047】ステップ220において、まず倒産確率を本システムによって計算するための元となる財務データなどの必要なデータが入力される。ステップ230において、ステップ220において入力されたデータに基づいて決定木が順次作成される。

【0048】ステップ240において対象企業のデータをステップ230において作成された決定木モデルに適用する。ステップ250において、決定木の診断結果から倒産確率が計算される。ステップ260において、本システムの一連の処理が終了する。

【0049】図3には、図2のステップ220において入力される財務データ300が例示されている。

【0050】図3における財務データ300は、各企業のID（識別番号）310、倒産（D）又は非倒産（N）を示すクラス（CLASS）320、売上高営業利益率（U r i R i e k i R）330、売上高税引前利益率（U r i Z M a e R i e k i R）340、売上高当期利益率（U r i T R i e k i R）350、総資本税引前利益率（S o S i h o Z M a e R i e k i R）360、自己資本率（J S i h o R）370等のデータを含むのが一般的である。これらのデータは、例えば帝国データバンク社等から入手可能である。

【0051】図4には、入力された財務データから決定木を作成する手順が示されている（図2のステップ230に対応）。ステップ410で、決定木を作成する手順が開始される。

【0052】ステップ420において、全財務サンプル・データが取り込まれる。全財務サンプル・データは、例えばビュー422に示されるように、倒産企業（D）の数が829であり、非倒産企業（N）の数が829である（422）。

【0053】ステップ430において、ノードの番号を表すnに初期値0を代入する。ステップ440において、n番目のノードにおいてサンプル・データで倒産企業（D）と非倒産企業（N）を最も偏らせる分割ルールを発見する。例えば、分割ルールとして、エントロピーの最小化基準等が使用される。

【0054】この発見された分割ルールを対象サンプル・データに適用したのがビュー442である。ビュー442のX軸（横軸）及びY軸（縦軸）はそれぞれある財務データを表している。

【0055】ビュー442において、445は倒産企業の割合が80～100%、446は倒産企業の割合が60～79%、447は倒産企業の割合が40～59%、448は倒産企業の割合が20～39%、449は倒産企業の割合が0～19%をそれぞれ表している。

【0056】例えば、横軸に自己資本率、縦軸に有利子負債金利をとると、右側の明るい部分が比較的安全な企業が多い領域になる。つまり、このルールは、「自己資本率の高い企業は、安定だが、低い企業については、有利子負債金利の高いところから危険になっていく」という傾向を示している。

【0057】すなわち、線444で囲まれた部分内の企業が比較的安定であり、線444で囲まれた部分外の企業が比較的危険であることが分かる。ステップ450において、ステップ440において発見されたルールを用いて対象データが2つに分割される。

【0058】現在のノードn452のすべてのデータが、領域の内側にあるデータ454と領域の外側にあるデータ456に分割され、決定木が作成される。ノード454に属するデータの内容は非倒産(N)が多く倒産(D)が少ないものとなる(455)。

【0059】これに対して、ノード456に属するデータの内容は倒産(D)が多く非倒産(N)が少ないものとなる(457)。ステップ460において、分割を更に続けるかどうか判断される。判断の基準の1つは、あるノードnにおけるデータの数(D+N)が所定の数(T)より小さいか否かで判断される。

【0060】ステップ460において、 $N+D < T$ を満たすならば一連の処理は終了し(ステップ470)、 $N+D < T$ を満たさないならばステップ480へ移行する。この例では、再帰的な処理の終了条件として「 $N+D < T$ 」を用いているが、本発明は他の条件に限定されるものではなく、適用対象に応じて他の条件を使用することも可能である。また、ポストプルーニング法やプリプルーニング法を用いても良い。

【0061】ステップ480において、ノードの番号nを1つ増分させて、次のノードの処理に移行する。ステップ460の条件を満足するまでステップ440～450における処理を再帰的に繰り返すことにより決定木が完成する。

【0062】図5には、図4の処理によって作成された実際の決定木が示されている。図5の決定木は、簡単のためノードの数を5つ(512～518)に限定している。

【0063】図5において、円で表したノードは中間のノードであり、正方形で表したノードはこれ以上分割されないノードを表している。ノード512がルートであり、ノード512における全サンプルデータは、非倒産(N)が829件であり、倒産(D)も829件である。

【0064】何らかの分割ルールを適用した結果、ノード514においては、非倒産(N)が627件であり、倒産(D)が124件と、圧倒的に非倒産データが多くなっている。すなわち、ノード512で適用した分割ルールが良いものであることを表している。ノード512の孫のノード518では、非倒産(N)が63件であり、倒産(D)が554件と、圧倒的に倒産データが多くなっている。すなわち、ノード516で適用した分割ルールが良いものであることを表している。

【0065】従って、対象となる企業は正方形のノード(514, 517, 518)のいずれかに分類されることになる。

【0066】図6は、診断対象企業の作成された決定木モデルへの適用手順が記載されている。これは、図2のステップ240に対応している。

【0067】ステップ620において、対象企業をデータベースから選択するか、又は対象企業のデータを新たに入手する。ステップ630において、図5の処理で作成した決定木モデルに診断対象企業を当てはめる。

【0068】図7は、図6の処理の結果が示されている。図5の処理により作成された決定木がビュー710の中に示されている。この場合は、決定木は、ルートを含めて5つのノード740～748を持っている。

【0069】いま診断対象としている企業1社は、決定木のノード740に分類されている。このノード740に分類された企業のデータ数は、非倒産(N)が627個で、倒産が(D)が124個であり、かなり非倒産企業の割合が高いことが分かる(720)。

【0070】企業1社は、ビュー730では、太線で囲まれた部分732に属する。すなわち、1社の財務内容は、安全な領域744の中である。従来の技術では、1社の倒産確率まで計算することができないため、1社の属する決定木のノード740に分類された企業のデータ、非倒産(N)が627個で、倒産が(D)が124個から定性的に倒産・非倒産を判断するだけであった。

【0071】図8は、決定木の診断結果から企業の倒産確率を計算する手順が示されている。これは、図2のステップ250に対応している。

【0072】ブロック810において、図7等を用いて説明した手順により求めた情報が事後確率を求めるパラメータとしてブロック840に提供される。一方、ブロック820において景気を表すデータ、例えば「日経平均株価」が入力される。

【0073】ブロック830において、景気を表すデータを用いて現在の平均的倒産確率が計算される。これは、例えば、 $A \times \text{日経平均株価} + B$ のような計算式で算出される。ブロック830で計算された現在の平均的倒産確率は、事前確率として、ブロック840に提供される。実施例では事前確率の計算法として「 $A \times \text{日経平均株価} + B$ 」を用いているが、事前確率の計算法として

他の計算式を用いることも可能である。

【0074】ブロック840は、事後確率を求めるパラメータとして決定木から提供される情報及び、事前確率としての平均的倒産確率の情報に基づき、ベイズの定理を用いて最終的に対象となる企業の倒産確率を定量的に*

ベイズの公式(式A)

A及びC、 C_1, C_2, \dots, C_n は事象であり、

$$C_1 \cup C_2 \cup \dots \cup C_n = \Omega$$

$$C_i \cap C_j = \emptyset, i \neq j$$

を満たすとする。このとき、

$$P(C_i | A) = \frac{P(A | C_i) P(C_i)}{P(A | C_1) P(C_1) + P(A | C_2) P(C_2) + \dots + P(A | C_n) P(C_n)}$$

($i = 1, 2, \dots, n$)

が成立する。

【0076】図9は、上述のベイズの公式を図5、図7及び図8で説明した事例に実際に適用した式である。図5、図7及び図8で説明した事例の場合、全サンプルデータの倒産企業数(D)は829個、全サンプルデータの非倒産企業数(N)は829個である。また、図7で説明した事例の場合、対象企業1社が分類されたノード740における倒産企業数(D)は124個、非倒産企業数(N)は627個である。

【0077】また、図8において説明した平均的倒産確率を1%と仮定する。図9における計算式900に上述の値を当てはめる。平均的倒産確率910、940、970には、1%、すなわち0.01が入力される。あるグループの倒産数920、950には、124が入力され、あるグループの非倒産数980には、627が入力される。

【0078】全サンプルデータの倒産数930、960には、829が入力され、全サンプルデータの非倒産数990には、829が入力される。以上の値を式900に代入し実際に倒産確率を計算すると、 $P=0.1994\%$ と求まる。従って、いま対象としている企業1社の倒産確率は0.1994%と求まった訳である。

【0079】図10には、本願発明を適用したシステムが例示されている。ユーザが、本システムを立ち上げた後に入力画面1010が表示される。ユーザは、企業コードや企業名(1030)や条件検索(1020)により、診断企業の選択を行い(1060)、日経平均株価を入力する(1080)とする。ユーザは、その他の必要な入力を済ませた後、診断ボタンをクリックすることにより診断を開始する(1070)。

【0080】図11には、本願発明を適用したシステムが例示されている。ユーザが図10において、開始した診断の診断結果が画面1100に示されている。定性的な診断結果である倒産判別がモデル年数とともに各モデルが示されている。決定木による判別は一番右に示されており(1120)、この場合は、すべて安全と判別されている。

* 計算する。

【0075】ここで、ベイズの定理とは、数理統計学者のベイズ(1706-1761)により発見された定理である。

【数1】

【0081】更に、定量的な判別が、倒産確率として示されている。決定木による倒産確率は、この場合、各モデル年で0.2%、0.3%及び0.4%と示されている(1130)。また、診断結果に用いられた財務データ等が同時に示されても良い(1140)。

【0082】以上の説明においては、企業の倒産確率について具体的な対象として説明したが、本願発明は、上述の企業の倒産確率の計算への適用に限定されるものではなく、複数の属性データを持つ過去のデータから将来のある事象の発生を計算する場合に広く適用可能である。例えば、自動車事故の発生確率の計算や、火災の発生確率、地震の発生確率、或いは、人の死亡確率、余命の計算などにも適用可能である。

【0083】

【発明の効果】本発明によって、学習データ内の倒産/非倒産の比率によらず現実的な確率を計算することができ、倒産確率の計算方法を提供することが可能となる。

【0084】また、本発明によって、学習データ内の倒産/非倒産の比率によらず現実的な確率を計算することができ、かつ、計算時点の景気状況を確率に反映させることができる倒産確率の計算方法を提供することが可能となる。

【0085】以下まとめとして他の実施例を記載する。

【0086】(1) 企業の債務不履行の発生確率を計算する方法であって、前記方法は、(a)複数の企業の財務データを入力するステップと、(b)前記財務データから決定木を作成するステップと、(c)対象企業の財務データを決定木に適用するステップと、(d)前記対象企業を決定木に適用した結果から債務不履行の発生確率を計算するステップと、を含む方法。

【0087】(2) 企業の債務不履行の発生確率を計算する方法であって、前記方法は、(a)複数の企業の財務データを入力するステップと、(b)前記財務データから決定木を作成するステップと、(c)対象企業の財務データを決定木に適用するステップと、(d)景気を表すデータから平均的倒産率を計算するステップと、

(e) 前記対象企業を決定木に適用した結果及び前記平均的倒産率から債務不履行の発生確率を計算するステップと、を含む方法。

【0088】(3) 企業の債務不履行の発生確率を計算するプログラムを記録した記録媒体であって、前記プログラムは、(a) 複数の企業の財務データを入力するステップと、(b) 前記財務データから決定木を作成するステップと、(c) 対象企業の財務データを決定木に適用するステップと、(d) 前記対象企業を決定木に適用した結果から債務不履行の発生確率を計算するステップと、を含むプログラムを記録した記録媒体。

【0089】(4) 企業の債務不履行の発生確率を計算するプログラムを記録した記録媒体であって、前記プログラムは、(a) 複数の企業の財務データを入力するステップと、(b) 前記財務データから決定木を作成するステップと、(c) 対象企業の財務データを決定木に適用するステップと、(d) 景気を表すデータから平均的倒産率を計算するステップと、(e) 前記対象企業を決定木に適用した結果及び前記平均的倒産率から債務不履行の発生確率を計算するステップと、を含むプログラムを記録した記録媒体。

【0090】(5) 前記(d)ステップが、ベイズの定理を利用することと特徴とする(1)乃至(4)に記載の方法。

【0091】(6) 企業の債務不履行の発生確率を計算する装置であって、前記装置は、(a) 複数の企業の財務データを入力する手段と、(b) 前記財務データから決定木を作成する手段と、(c) 対象企業の財務データを決定木に適用する手段と、(d) 前記対象企業を決定木に適用した結果から債務不履行の発生確率を計算する手段と、を含む装置。

【0092】(7) 企業の債務不履行の発生確率を計算する装置であって、前記装置は、(a) 複数の企業の財務データを入力する手段と、(b) 前記財務データから決定木を作成する手段と、(c) 対象企業の財務データを決定木に適用する手段と、(d) 景気を表すデータから平均的倒産率を計算する手段と、(e) 前記対象企業を決定木に適用した結果及び前記平均的倒産率から債務不履行の発生確率を計算する手段と、を含む装置。

【0093】(8) 企業の債務不履行の発生確率を計算するシステムであって、前記システムは、プロセッサ、記憶装置、入力装置、表示装置を含み、更に、(a) 複数の企業の財務データを入力する手段と、(b) 前記財務データから決定木を作成する手段と、(c) 対象企業の財務データを決定木に適用する手段と、(d) 前記対象企業を決定木に適用した結果から債務不履行の発生確率を計算する手段と、を含むシステム。

【0094】(9) 企業の債務不履行の発生確率を計算するシステムであって、前記システムは、プロセッサ、記憶装置、入力装置、表示装置を含み、更に、(a) 複

数の企業の財務データを入力する手段と、(b) 前記財務データから決定木を作成する手段と、(c) 対象企業の財務データを決定木に適用する手段と、(d) 景気を表すデータから平均的倒産率を計算する手段と、(e) 前記対象企業を決定木に適用した結果及び前記平均的倒産率から債務不履行の発生確率を計算する手段と、を含むシステム。

【0095】(10) 前記(d)手段が、ベイズの定理を利用することと特徴とする(6)乃至(9)に記載の方法。

【0096】(11) 前記財務データが、倒産企業のデータ及び非倒産企業のデータを含むことを特徴とする(1)乃至(10)に記載の方法。

【0097】(12) 自動車事故の発生確率を計算する方法であって、前記方法は、(a) 複数の自動車事故データを入力するステップと、(b) 前記自動車事故データから決定木を作成するステップと、(c) 対象となる自動車事故データを決定木に適用するステップと、(d) 前記対象となる自動車事故データを決定木に適用した結果から対象となる自動車事故の発生確率を計算するステップと、を含む方法。

【0098】(13) 火災の発生確率を計算する方法であって、前記方法は、(a) 複数の火災データを入力するステップと、(b) 前記火災データから決定木を作成するステップと、(c) 対象となる火災データを決定木に適用するステップと、(d) 前記対象となる火災データを決定木に適用した結果から対象となる火災の発生確率を計算するステップと、を含む方法。

【0099】(14) 企業の債務不履行の発生確率を計算する方法であって、前記方法は、(a) 複数の企業の財務データを入力するステップと、(b) 前記財務データから決定木を作成するステップと、(c) 対象企業の財務データを決定木に適用するステップと、(d) 前記対象企業を決定木に適用した結果から債務不履行の発生確率を計算するステップを含み、前記(c)ステップが、前記複数の企業を倒産企業が多いグループと非倒産企業が多いグループの2つのグループに分類するためのデータを前記財務データの中から発見するステップを含む方法。

【0100】(15) 人の死亡確率を計算する方法であって、前記方法は、(a) 複数の人のデータを入力するステップと、(b) 前記人のデータから決定木を作成するステップと、(c) 対象となる人のデータを決定木に適用するステップと、(d) 前記対象となる人のデータを決定木に適用した結果から対象となる人の死亡確率を計算するステップと、を含む方法。

【0101】(16) 地震の発生確率を計算する方法であって、前記方法は、(a) 複数の地震のデータを入力するステップと、(b) 前記地震のデータから決定木を作成するステップと、(c) 対象となる地震のデータを

決定木に適用するステップと、(d)前記対象となる地震のデータを決定木に適用した結果から対象となる地震の発生確率を計算するステップと、を含む方法。

【0102】(17)前記(d)ステップが、ベイズの定理を利用することを特徴とする(12)乃至(16)に記載の方法。

【図面の簡単な説明】

【図1】図1には、本発明を適用した企業の倒産確率計算プログラムの機能的構成が示されている。

【図2】図2は、本発明を適用したシステムの主な処理の流れを示したフローチャートである。

【図3】図3には、財務データ300が例示されている。

【図4】図4には、入力された財務データから決定木を作成する手順が示されている。

【図5】図5には、図4の処理によって作成された実際の決定木が示されている。

【図6】図6は、診断対象企業を作成された決定木モデ

*ルへの適用手順が記載されている。

【図7】図7は、図6の処理の結果が示されている。

【図8】図8は、決定木の診断結果から企業の倒産確率を計算する手順が示されている。

【図9】図9は、ベイズの公式を実際の事例に適用した式が示されている。

【図10】図10には、本願発明を適用したシステムが例示されている。

【図11】図11には、本願発明を適用したシステムが例示されている。

【符合の説明】

110 企業データの入力・選択部

120 倒産確率を計算するためのモデル部

130 景気を表すマクロ情報の入力部

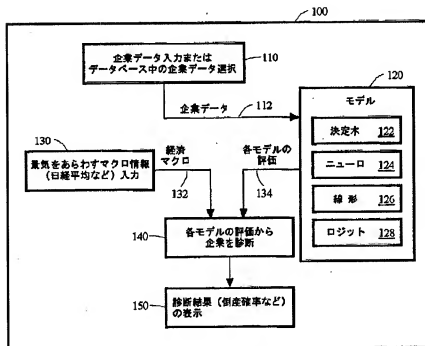
140 各モデルの評価から企業を診断する診断部

150 診断結果の表示部

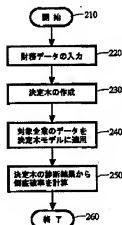
1010 入力画面

1100 診断結果の表示画面

【図1】



【図2】



【図3】

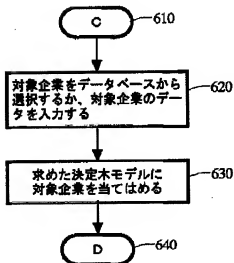
300

0817	N	-0.063776765	-0.063776765	-1.382809507	-2.143293584	40.5411304
0818	N	4.8241084	2.8277039	1.346468335	1.953183197	65.6240130
0819	N	2.493991974	1.079859379	0.396952377	1.203978041	3.28509461
0820	N	2.440831204	2.728008161	1.383007025	5.036844857	27.1731668
0821	N	-0.961880528	-0.543400039	-0.543400039	-0.805639557	77.2731668
0822	N	4.870755442	1.381302323	1.107718363	-0.805639557	7.23989965
0823	N	0.100398179	-5.867208472	-5.879069406	1.189247492	3.11408066
0824	N	0.271207078	-0.603199351	-0.603199351	-4.426556304	10.6222047
0825	N	1.381471653	1.40195923	0.823291964	-0.979933068	15.9912513
0826	N	-4.103604414	-6.371044793	-6.371044793	4.795538648	21.9298656
0827	N	11.32458659	16.01090951	9.407785944	-9.804108108	31.4702202
0828	N	4.303669107	2.125716324	0.904190491	18.78208695	12.6039584
0829	N	3.7300625007	2.87373511	1.238202129	2.075266926	7.26102459
0830	D	2.188901843	0.275794318	0.109121605	3.821820084	3.79563757
0831	D	0.283974265	0.561420387	0.389627123	0.391235926	14.8604930
0832	D	2.177354711	0.27855381	0.27855381	0.695387208	4.41951612
0833	D	0.548498876	0.14478052	0.144780595	0.376297782	10.2474360
0834	D	3.445800385	0.394964145	0.211460764	0.193645588	3.71408170
0835	D	0.804289544	0.268096515	0.144772118	0.437980831	13.7247292
0836	D	9.195430056	-3.493867379	-3.49386739	0.357781753	34.935753
0837	D	0.013518976	-4.088505982	-4.088585982	-3.795074377	-2.4008867
0838	D	2.612621845	0.141739332	0.074577025	0.151317518	2.83049533

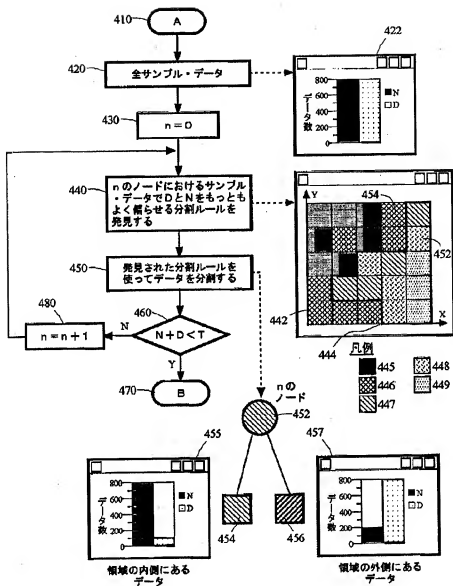
310 320 330 340 350 360 370

310 ... ID: 企業の名
 320 ... CLASS: Nが非倒産、Dが倒産を示す
 330 ... UnRiskR: 売上高営業利益率
 340 ... UnZMacRiskR: 売上高役員前利益率
 350 ... UnTRiskR: 売上高当期利益率
 360 ... SoShoMacRiskR: 総資本税引前利益率
 370 ... JSR: 自己資本比率

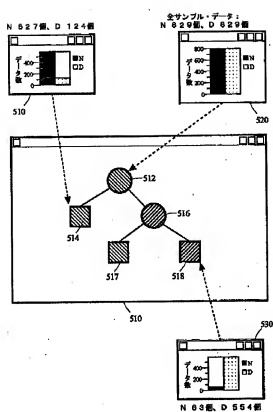
【図6】



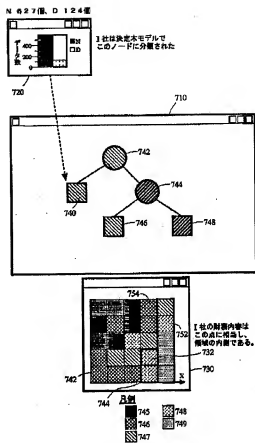
【図4】



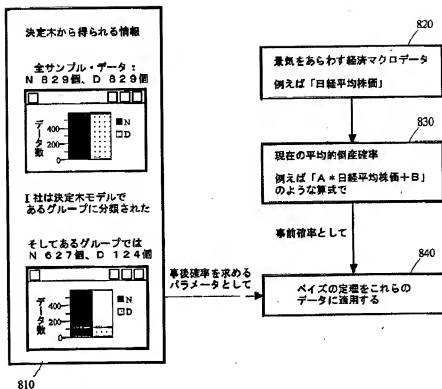
【図5】



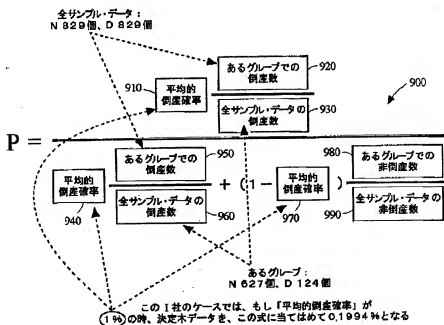
【図7】



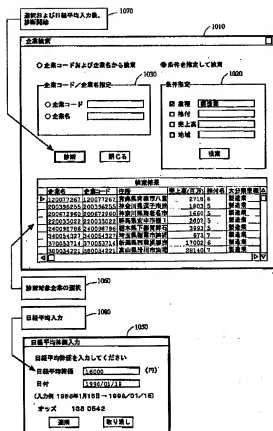
【図8】



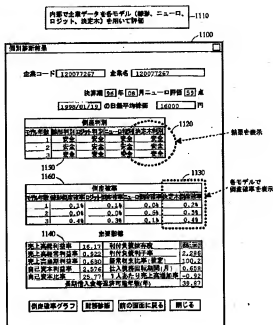
【図9】



【图 10】



【図 1-1】



フロントページの続き

(72)発明者 森本 康彦
神奈川県大和市下鶴間1623番地14 日本ア
イ・ピー・エム株式会社 東京基礎研究所
内

(72)発明者 熊田 清志
神奈川県大和市下鶴間1623番地14 日本ア
イ・ピー・エム株式会社 東京基礎研究所
内

(72)発明者 福田 剛志
神奈川県大和市下鶴間1623番地14 日本ア
イ・ピー・エム株式会社 東京基礎研究所
内

F ターム(参考) 5B049 BB46 EE03 EE05 EE14 EE24
EE31 FF03 FF04